Response to Reviewer 1 "Snow Depth Estimation on Lead-less Landfast ice using Cryo2Ice satellite observations" by Saha et al.

Summary:

The study assesses the potential for near-coincident ICESat-2 and Cryosat-2 (Cryo2Ice) satellite data in estimating snow depth over landfast ice in the Canadian Arctic Archipelago. Snow depth is retrieved by calculating the absolute difference in surface height from the two satellites, considering an ocean tide correction. The study compares the retrieved snow depths from Cryo2Ice with in-situ measurements, showing good agreement in terms of mean values. However, Cryo2Ice snow depths were, on average, underestimated by 20.7%. Discrepancies are attributed to differences in sampling resolutions, snow characteristics, surface roughness, and tidal correction errors. The results suggest the potential for estimating snow depth over lead-less landfast sea ice, but further investigation is needed to understand biases related to sampling resolution, snow salinity, density, surface roughness, and altimeter correction errors.

The authors thank the reviewer for the very comprehensive and accurate overview of the study.

General Comments:

This is an interesting study, which will be valuable to improve our understanding with respect to retrieving snow depth from a dual-altimeter approach. I had no problems to follow the paper, but I believe clarity can be improved. However, there are some parts in the analysis and discussion, which I think need some clarification and revision. I think this work deserves publication, but major revisions are needed.

My main concerns are:

- There is quite some focus on the tidal correction, which I agree is important. But one of the main limitations, from my point of view, is the **large CS2 footprint and the noise in the CS2 height retrievals**, which is not surprising as we know from previous studies. But considering the relatively small sample size, it will have a large impact. The reasons for the CS2 height uncertainties are discussed by the authors, e.g., the surface roughness, snow salinity, scattering horizons in the snow layer. And given the significant difference in footprint between IS2 and CS2, we can hardly assume that CS2 heights will represent the corresponding snow-ice interface, even if colocated. Considering the snow depth distribution from in-situ measurement sites indicates how much spread is just within one CS2 footprint. In addition, the retracker used for the CS2 height retrievals is a threshold retracker, using a fixed threshold, where we cannot be sure if it is tuned exactly for the same ice conditions we find in this area. I believe this study can contribute to characterize and **quantify uncertainties and limitations of this approach**, but it **should be emphasized more clearly**. For example, I think the comparison between the mean/median values of Cryo2Ice and in-situ measurements has only limited meaning, which brings me to my next point.

The authors thank the reviewer for their valuable observations on the analysis. The authors agree that the difference in CS2 and IS2 footprint is definitely one of the major challenges of getting coincident snow depths from Cryo2Ice. Therefore, in order to reduce the noise, a smoothening of CS2 footprints are often suggested. However, given the fact this study considers snow depth over ~75 kilometers and *tries to identify the uncertainties in combining CS2 and IS2 over landfast sea ice, a smoothening approach may lead to loss of distinctive roughness features and also a reduction in overall resolution of the snow depth retrieved. Averaging IS2 over ~300 metre along-track resolution allows us to make the closest comparison to the 300 meter along-track footprint of CS2 and ~200 meter in-situ snow depth retrievals.*

We notice that smoothing the ATL07 data over ~1km makes us loose some of the distributive ridges which IS2 is very sensitive to (Figure R1). Therefore, in order to understand the impact of roughness features on the overall snow depth retrievals, the authors believe averaging IS2 strong beam heights over ~300 meter (across track resolution of CS2) captures enough of the variability while maintaining the highest resolution. The impact of the difference in footprint size on the retrieved snow depth is eventually addressed when we compare the Site specific snow depth distributions (Section 4.2). The authors agree with the reviewer that further emphasis on the quantification of the uncertainties need to be added to the discussions and shall be included in the revised paper. We do agree that some further statistical analysis comparing the in-situ and Cryo2Ice distributions need to be included in the revised paper.

Therefore, we include (a) further comparison of the shapes of snow depth distributions including comparing the skeness and and kurtosis of the site wise in-situ vs Cryo2ice snow depth values and (b) comparison of the site-wise standard deviations of in-situ vs Cryo2Ice (c) correlating IS2 and CS2 ellipsoidal heights with the retrieved snow depths.

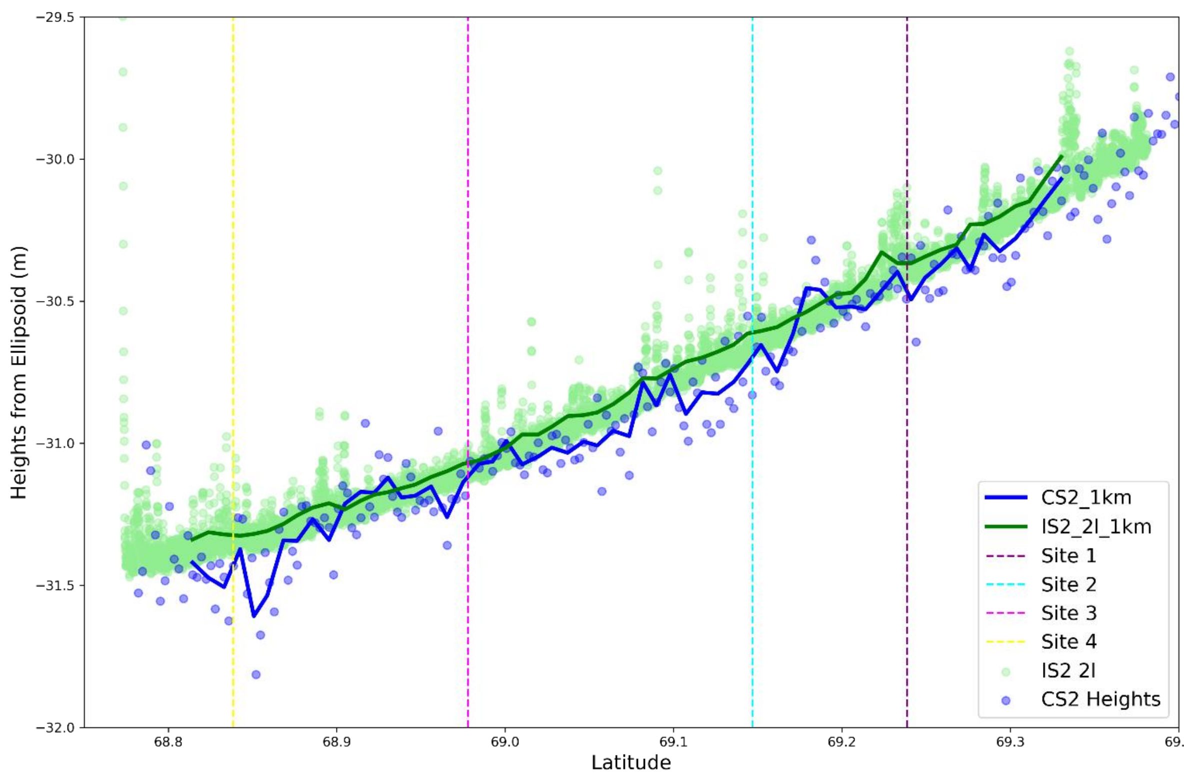

Figure R1: IS2 and CS2 Height distributions. IS2 2l height in native distribution is shown in green while CS2 height is shown in purple. The blue line shows the smoothed IS2 and CS2 lines over 1-km.

- I am not an expert in statistics, but some of the decisions made in the processing need some clarification and potential revision. I do not think it is a good idea to **just drop negative snow depth values.** From a physical point of view, they do not make sense, but statistically they are important, **reflecting the impact of uncertainties**. By removing only negative values, your snow depth retrieval is likely biased towards higher values. The negative values should be **also part of all the histograms** because that's the reality when you subtract one height retrieval from another, when both come with uncertainties. Moreover, some steps and figures need to be explained in more detail, see therefore the specific comments below.

  The authors agree with the reviewer that it is important to include the negative snow depth values as part of the final computation of the snow depths. However, the impact of large negative snow depths caused largely due to the noise of CS2 needs to be discarded in the final snow depth estimation. As shown in Figure R2, there is a significant portion of negative snow depths lower than -10 cm which also corresponds with noise portions of the CS2 ellipsoid height which ultimately biases low the mean snow depth retrievals. Therefore, we disregard negative snow depths that are lower than 2 standard deviations from the mean and attribute this to the uncertainty due to the CS2 noise. Therefore, after adjusting for the outliers in the negative snow depths, we get a mean snow depth distribution of 7.52 cm. The new distribution of the Cryo2Ice snow depths are presented in Figure R3.

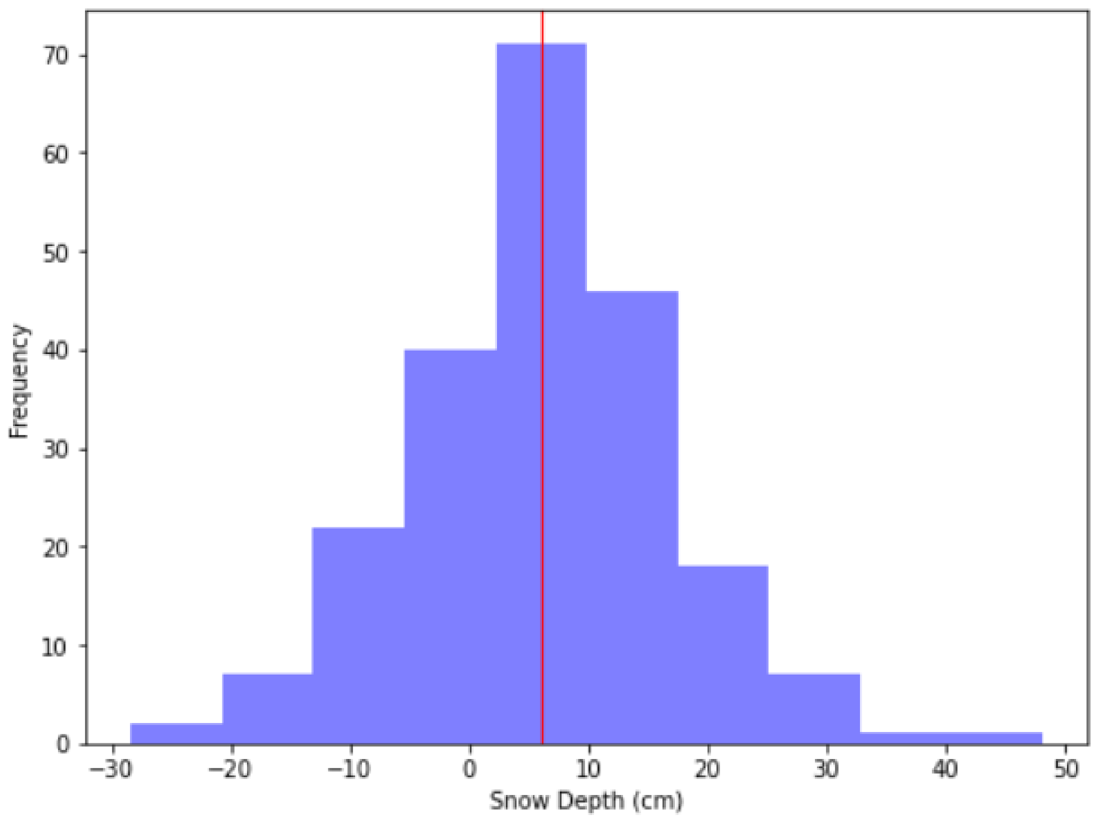

Figure R2: Distribution of raw Cryo2Ice snow depth values including negative snow depth values

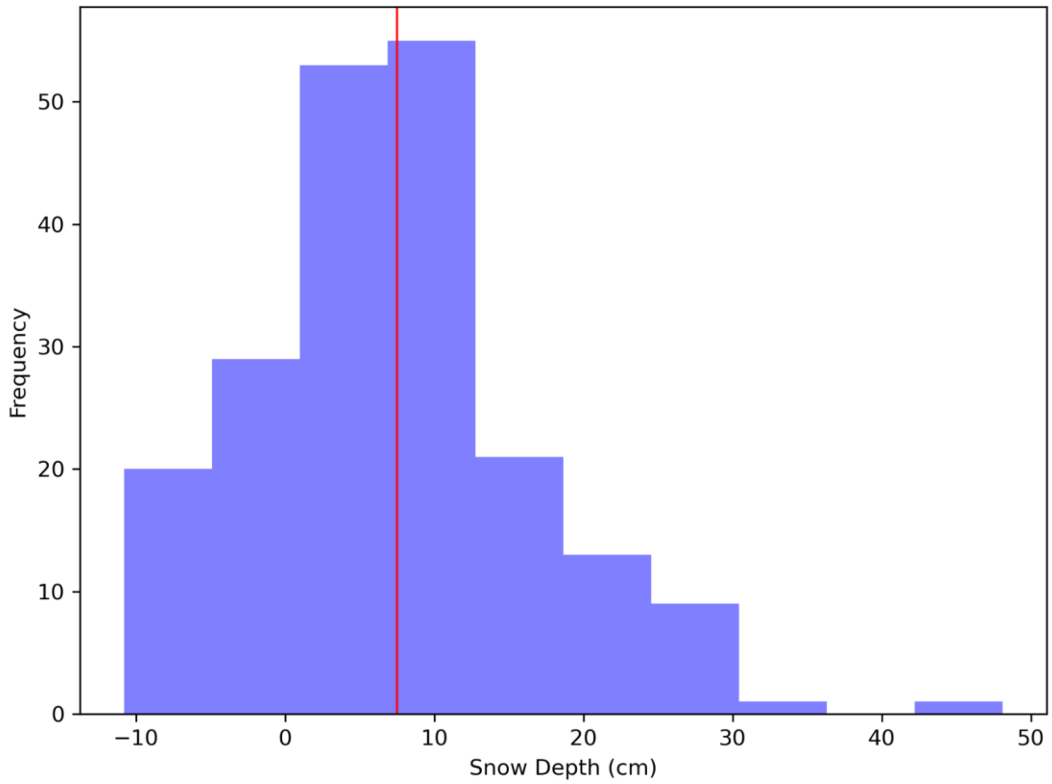

*Figure R3: Cryo2Ice Snow Depths adjusted for negative snow depth values greater than 2 standard deviation from the mean snow dept*h

Specific Comments:

L101: Figure 3 is introduced before Figure 2.

This was a mistake on part of the author. The Figure number has been changed to Figure 1 instead of Figure 3.

L113: For the snow depth measurements, what was your sampling strategy? Can you be more specific here? Did you walk straight transects? Did you ensure representative sampling, considering the fraction of deformed sea ice?

The authors thank the reviewers for their question regarding the sampling strategy. The authors agree that further details about the sampling strategy is warranted in the paper. The transects were set considering wind direction as well as the sea ice surface features for each spot. As shown in Figure R4 (which will be included in the appendix of the revised paper), the shape of the transects are demonstrated along with the representative snow depths for each site. The sampling strategy was to ensure that we cover Cryo2Ice along-track and across-track directions, wind direction and sample *the different representative roughness features. In Site 1, two L-shaped transects representing the rough and smooth sea ice zones were conducted (Figure R4 (a). For Site 2, which had a large ridge which covered a significant portion, we tried to take two different L-shaped transects crossing the ridge at least four times but ensuring we get the near-ridge features as well as the smoother zones further away from the ridges (Figure R4 (b)). For Site 3 and 4 which had wider regions of smooth and rough sea ice respectively, two L-shaped transects were conducted (Figure R4 (c) & (d)).*

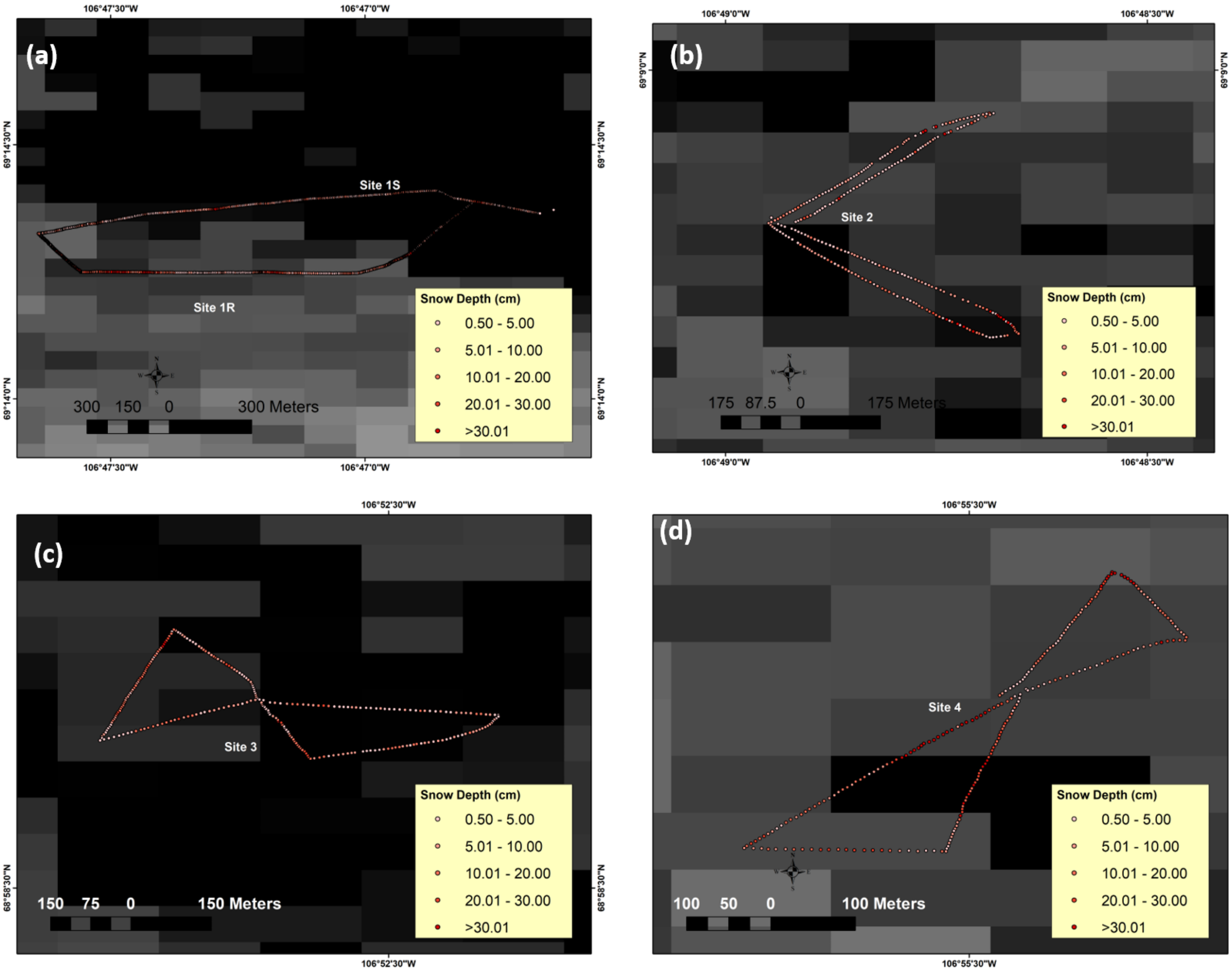

Figure R4: The in-situ sampling sites showing the position of the magnaprobe samples for (a) Site 1R and 1S, (b) Site 2, (c) Sites 3 and (d) Site 4.

L157: The MSS is not mentioned under 2.6. I suggest to briefly explain the reason here.

The line now reads: "However, the mean sea surface (MSS) ensuring that the IS2 ATL07 heights are referenced to the WGS84 ellipsoid. The MSS is calculated based on decadal averages and therefore *are not representative of the variation of sea surface heights within the 77 minutes' interval between the IS2 and CS2 passes".*

L164: To my knowledge, the ATL07 product does not contain the individual photon heights, but segments of different length that aggregate the photon heights from the ATL03 product. I assume you have used these segments?

Yes, we have used the aggregated photon heights used in the ATL07 product. On average each segment i.e. difference between two aggregated ATL07 heights is 8.3 meters. We plan to include the following statement in L71 "ATL07 Heights are aggregated from ATL03 photon heights over variable distances, the heights were aggregated over 8.3 meters on average over the portion of the track used in this study."

L167: Are the retrieved Cryo2Ice snow depths not arranged along a straight line? Why then investigating spatial autocorrelation? Isn't it nearly 1D? Moreover, when I look at Fig. 8 (bottom plot), I find it hard to imagine how this works. The sample size is not very high and there is a lot of noise. And the spacing between point is already 300 m. Can you show a variogram? (Just in the response, does not need to go into the manuscript).

The authors agree that only the along-track variability in snow depth is being represented in the variogram given that the snow depths are colocated to the CS2 POCA points. However, the variogram gives us a sense of the spatial autocorrelation along-track which is useful to decide on which lag distance the snow depth points are auto correlated. Therefore, we observe that the variogram in Figure R4 starts to reach a inflection point around 300 meter and reaches a sill ~ 1km. We also notice that beyond the 300 meter lag distance, the semi-variance starts to become constant. Due to the high variability in the data, the variogram analysis doesn't provide a concrete lag distance but does give an indication that beyond the 1 km point, the spatial autocorrelation would be negligible. Therefore, while we don't believe the variogram gives a concrete evidence of which lag distance to consider for averaging, we can say that beyond the 1km distance, the snow depth points are not autocorrelated.

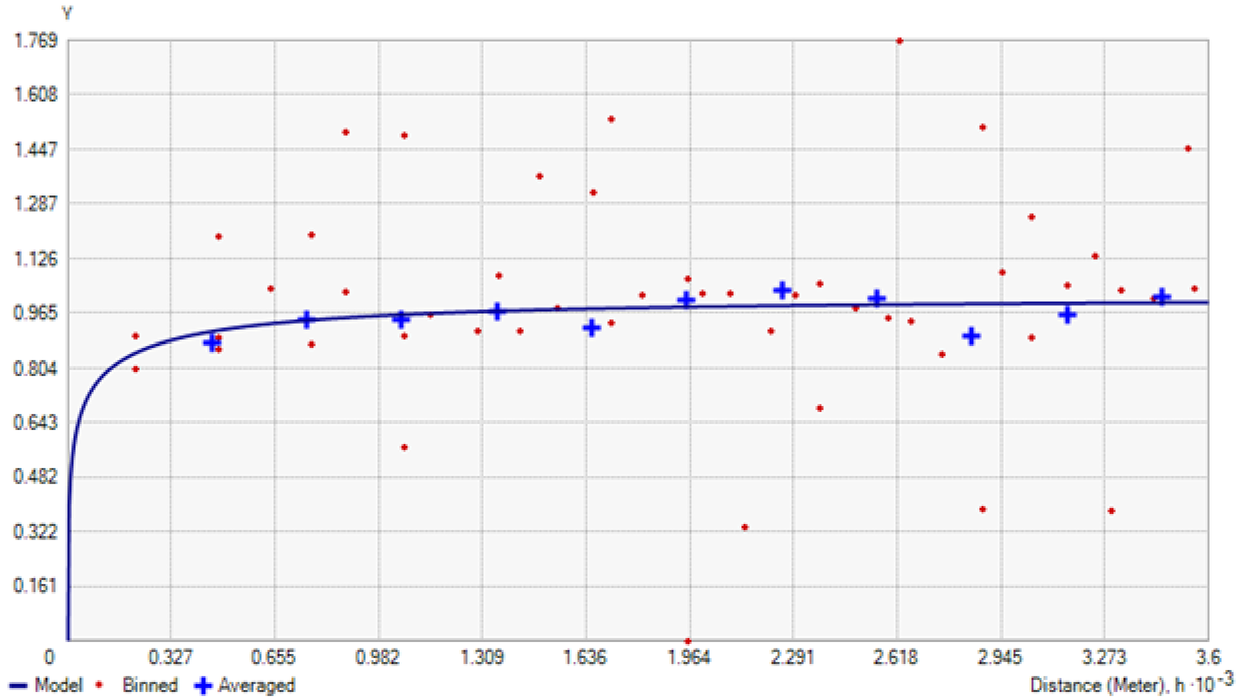

Figure R5: Semi-variogram analysis of the Cryo2Ice snow depths. Y axis represents the semi-variance while the x-axis shows the lag-distance.

L196: That's a nice approach with the Sentinel-1 backscatter. I suggest checking the "stability" and representativeness of the ICESat-2 heights, making use of the other beams. Just compare the height distributions from the 3 strong beams for the area of interest.

The authors thank the reviewer for their valuable suggestions. We compared the IS2 height distributions which are presented in the Figure R6. It is noteworthy that IS2 2l strong bema's height distribution is most similar to CS2 height distributions. This is because the distance of the strong beam is within ~1.5 kilometer of the CS2 Points of Closest Approach (POCA). Therefore, we believe using only the IS2 2l strong beam which is the closest to the POCA CS2 would make the best colocation of the IS2 and CS2 tracks. This is based on the assumption that the snow distribution corresponding to the CS2 POCA and the closest IS2 strong beam would be most similar. This assumption is further tested by comparing the Sentinel-1 Backscatter retrieved from all the IS2 strong beams and CS2 (Figure R7). We notice that while the mean ellipsoidal heights are similar, there is significant differences among the distributions of both IS2 1l and IS2 3l strong beams. Therefore, we don't consider the IS2 1l and IS2 3l strong beams for the subsequent colocation and snow depth retrieval steps.

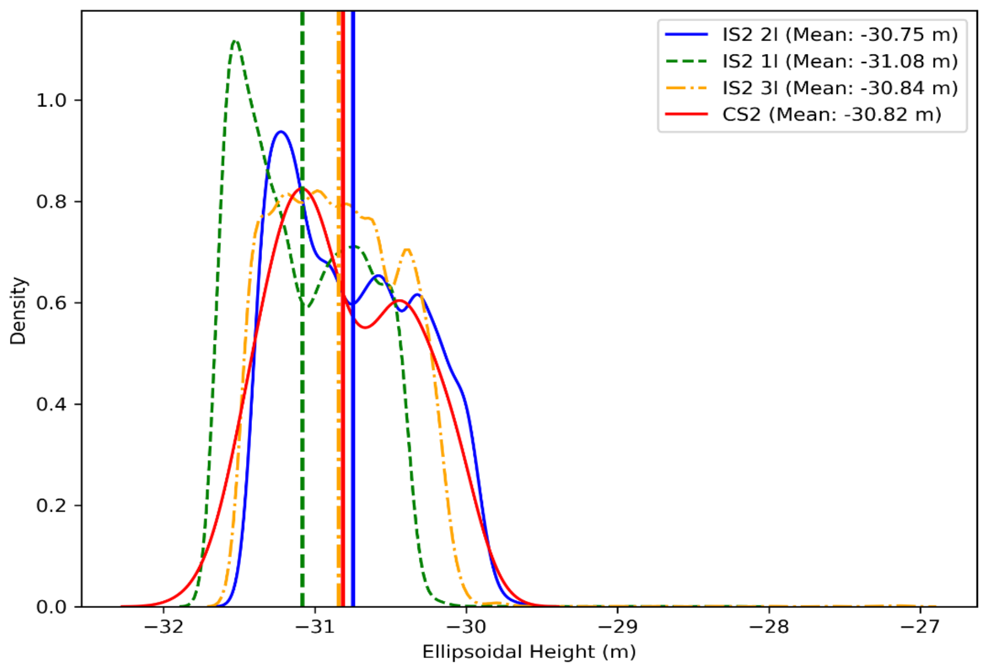

Figure R6: IS2 strong beam height distributions

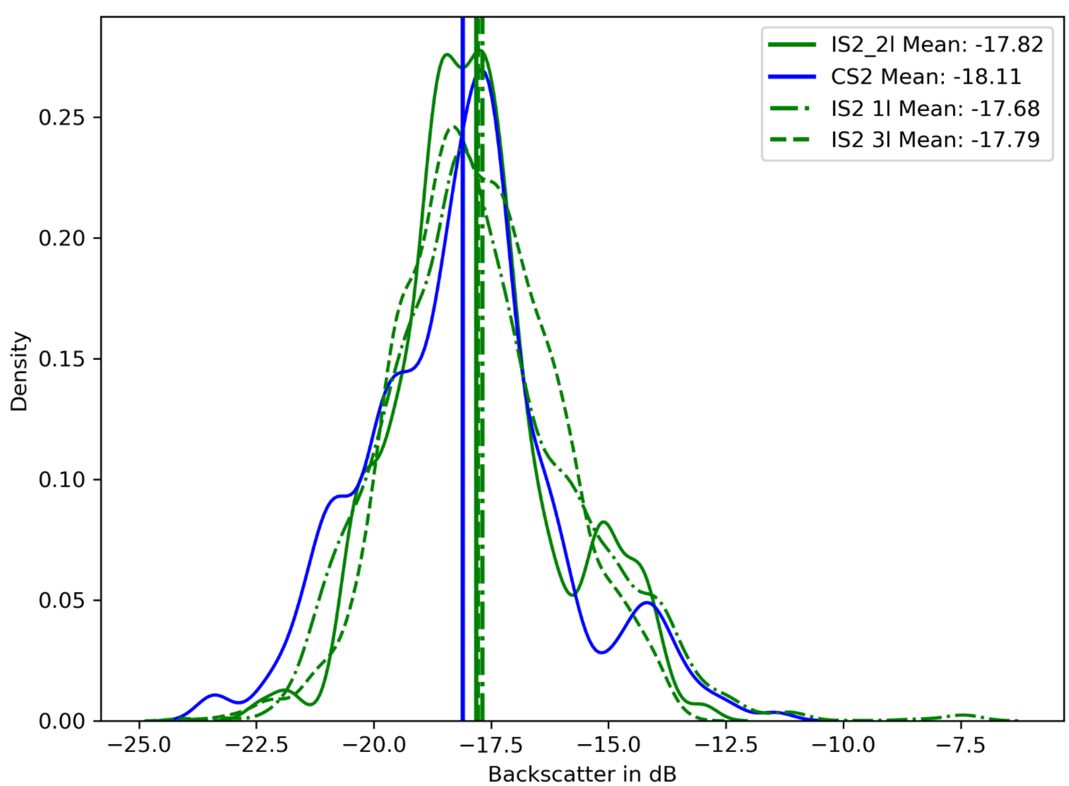

Figure R7 Comparing the backscatter retrieved from all IS2 strong beams with CS2

Figure 4: I suggest changing the legend. It is misleading. It looks like the backscatter of IS2/CS2 is shown here…

Noted. We changed the legend and the caption will also elaborate on this.

L206: Is this related to Figure 11? May be show this together with Figure 4? Farrell et al. (2020) primarily use ATL03. The ATL07 segments can be quite long. How many segments do you get on average within the 300 m segments? Can you derive a meaningful roughness from this?

L206 was intended to introduce surface roughness as a potential uncertainty impacting retrieved surface roughness. In order to compute surface roughness, Farrell et al., (2020) uses ATL07 segments but over 25 km long segments. Given our study area is ~75km long, this approach would not be able to different the difference in roughness zones. Therefore, we compute the roughness over ~300 metres instead. We get 36 ATL07 height segments within each 300 metre segments for the portion of the IS2 track that was studies. Therefore, each ATL07 segment is ~8.3-metre-long which is sufficient to portray the regional (~200 meters for each site) difference between each surface. The University of Maryland- Ridge Detection Algorithm (UMD-RDA) in Farrell et al., (2020) uses the ATL03 photon heights retains much finer resolution (0.7 metre) which is aimed towards ridge detection. However, we are more interested in representative roughness over ~200 metre regional sea ice roughness zones (Sites 1 to 4) which we believe is sufficiently captured by the 8.3 metre resolution of the ATL07.

L236: Figure 7 -> Figure 6?

The authors thank the reviewer for pointing out the incorrect Figure number. Figure number changed to Figure 6 instead of Figure 7.

Figure 5: The blue line is not explained.

The blue line is the Probability Density curve for the respective distributions. We will elaborate this in the caption.

L251: I don't see the negative values in Figure 8. I suggest adding a class with a specific colour for values <0. From Fig. 7, it does not look like negative values primarily occur close to the coasts.

The authors agree with the reviewer that the negative values need to be included in the figure and Figure 7 will be revised accordingly.

L252: I would argue that with removing negative values, you introduce a positive bias in the snow depth retrieval. It will only make sense if you assume that underlying uncertainties affect the snow depth exclusively in one direction. But looking at Figure 7, it just seems that there is significant noise on the CS2 heights, which goes in both directions (positive and negative).

The authors agree with the reviewer that the negative values need to be included in the subsequent snow depth calculations and the uncertainty calculations.

L295: I haven't fully understood why this test is done. "The test results show significant difference between in-situ sites which was also evident in the corresponding Cryo2Ice snow depths." Which are the corresponding Cryo2Ice snow depths? I guess there are just a handful in the vicinity of each site?

The Kruskal-Wallis non-parametric test was conducted to test whether the retrieved Cryo2Ice snow depths have similar site to site differences as obtained from the field. The test answers the question 'Site 2 has significantly different snow depths compared to Site 1 as seen from the field data, is this also true for the Cryo2Ice snow depths from Site 1 and 2?'. The test is important to check if the retrieved Cryo2Ice snow depths are realistic. The corresponding snow depths are the closest Cryo2Ice snow depth points which are within a similar roughness zone (Figure 8- Yellow lines). The roughness zones are defined for each site as the length of the Cryo2Ice track that has roughness (calculated from IS2) within 1 standard deviation of that obtained from the Site. Each Site has different number of Cryo2Ice snow depths that are within the same roughness zone.

L299: Related to the previous question: How many Cryo2Ice snow depths are you using for the comparison?

The number of Cyo2Ice snow depths varied based on the definition of the similar roughness zones (Figure 8, Yellow Lines). That is, the Cryo2Ice snow depth points which fall within 1 standard roughness are considered corresponding to each site. The number of Cryo2Ice snow depth sites compared to each site is presented in the Table below:

Table R1: The number of Cryo2Ice snow depths are within the similar roughness zones for each site

Site	**Number of corresponding Cryo2Ice snow depths**
Site 1	11
Site 2	19
Site 3	30
Site 4	7

Figure 10: I suggest showing the "raw" distributions, not the density functions. Again, how many Cryo2Ice samples have been used at each site for the PDFs?

Given the difference in frequency between the number of In-situ sites and the Cryo2Ice, the visualization of the raw distributions is challenging and therefore the density functions are used. The number of Cryo2Ice samples will be included in the PDFs. Table R1 provides the number of Cryo2Ice snow depths used for each site.

Figure 7: It would be also interesting to see the IS2 heights from the co-registration, averaged on the 300 m segments. Perhaps you can add them here?

The IS2 300 meter segment heights have been included in Figure 7 as follows:

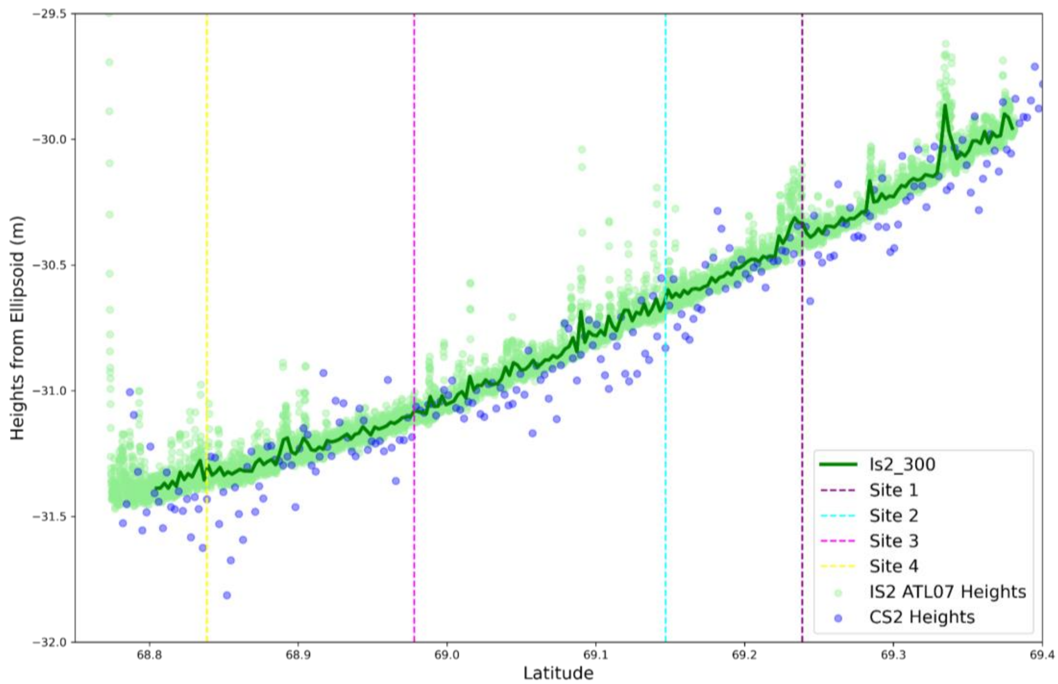

Figure 8: I suggest adding the mean and standard variation from the in-situ measurements at the 4 sites.

Noted. The mean and stand variations will be added to the to Figure 8.

L363: R**2 = 0.04 basically means no correlation I believe. But considering the noise level, especially from the CS2 heights, and the relatively low sample size, I wouldn't expect a higher R here.

The authors agree with the reviewer and the line is revised to:

*"There was no significant correlation (R**2 = 0.04) between IS2 and CS2 snow depths"*